

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

5-1980

Evaluation of Multivariate Homogenous Arma Model

Lucy Chienhua Tseng
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Applied Statistics Commons](#)

Recommended Citation

Tseng, Lucy Chienhua, "Evaluation of Multivariate Homogenous Arma Model" (1980). *All Graduate Theses and Dissertations*. 6945.

<https://digitalcommons.usu.edu/etd/6945>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



EVALUATION OF MULTIVARIATE HOMOGENEOUS ARMA MODEL

by

Lucy Chienhua Tseng

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1980

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to Dr. Ronald V. Canfield, my major professor, for his patient guidance and great help.

I would also like to thank Dr. Rex L. Hurst of my Graduate Committee for his helpful suggestions, and Professor David S. Bowles for his critical review of the thesis.

Lucy C. Tseng

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	v
ABSTRACT	vi
Chapter	
I. INTRODUCTION	1
II. CHARACTERIZATION OF MULTIVARIATE HOMOGENEOUS ARMA MODEL	5
Mathematical characterization	5
Physical characterization	6
III. MAXIMUM LIKELIHOOD ESTIMATION METHOD	8
Maximum likelihood estimation	9
Conditional maximum likelihood estimation	9
Unconditional maximum likelihood estimation	11
General procedure for calculating the unconditional sum of squares	12
Estimation of variance-covariance matrix	13
IV. O'CONNELL'S ESTIMATION METHOD	15
Method I	17
Method II	18
Bias correction	19
Agreement with Hurst Law	19
V. EVALUATION OF HOMOGENEOUS MODEL	22
Introduction	22
Expected values of correlation matrices	25
Results	28
Conclusions and recommendations	31
REFERENCES	34

TABLE OF CONTENTS (Continued)

	Page
APPENDICES	35
Appendix A. Computer Program in Maximum Likelihood Method	36
Appendix B. Computer Program in O'Connell's Method	45
Appendix C. O'Connell's Table	51

LIST OF TABLES

Table		Page
1.	Population model parameters C, E, and G and their lag 0,1,2 correlation matrices M_0, M_1, M_2	29
2.	Estimated parameters C, E, and G and their MSE to population parameters. Estimated lag 0,1,2 correlation matrices and their MSE to population correlations	30
3.	Estimated parameters C, E, and G and their MSE to population parameters. Estimated lag 0,1,2 correlation matrices and their MSE to population correlations by O'Connell's method	32

ABSTRACT

EVALUATION OF MULTIVARIATE HOMOGENEOUS ARMA MODEL

by

Lucy Chienhua Tseng, Master of Science

Utah State University, 1980

Major Professor: Dr. Ronald V. Canfield
Department: Applied Statistics

The purpose of this thesis is to study a restricted multivariate ARMA model, called the Homogeneous Model. This model is defined as one in which each univariate component of the multivariate model is of the same order in p and q as it is in the multivariate model.

From a mathematical respect, multivariate ARMA model is homogeneous if, and only if, its coefficient matrices are diagonal. From a physical respect, the present observation of a phenomenon can be modeled only by its own past observation and its present and past "errors."

The estimation procedures are developed based on maximum likelihood method and on O'Connell's method for univariate model.

The homogeneous model is evaluated by four types of data. Those data are generated reflecting different degrees of nonhomogeneity. It is found that the homogeneous model is sensitive to departures from the homogeneous assumptions. Small departures cause no serious problem, however, large departures are serious.

(54 pages)

CHAPTER I

INTRODUCTION

A mathematical model is often used to describe the behavior of a physical phenomenon. However, unknown and unpredictable factors do not permit deterministic models. Stochastic models are defined that can be used to predict future behavior as the probability of a future value lying between two specified limits.

Box and Jenkins (1970) first proposed the ARMA(p,q) time series stochastic model, which can be used to generate the values of some time-dependent quantity. The name ARMA(p,q) denotes a combination of autoregressive (AR) and moving average (MA) models. Autoregressive model is a way of expressing each observation in terms of past observations. The number p is an integer denoting the number of past observation or autoregressive terms in the model. The moving average model is a way of expressing each observation in terms of current and past disturbances. The value q is an integer denoting the number of past disturbances used in the model. Such models may be either univariate or multivariate. The ARMA(p,q) is expressed in this form:

$$\vec{Z}_t = C_1 \vec{Z}_{t-1} + C_2 \vec{Z}_{t-2} + \dots + C_p \vec{Z}_{t-p} + \vec{\epsilon}_t - E_1 \vec{\epsilon}_{t-1} - E_2 \vec{\epsilon}_{t-2} - \dots - E_q \vec{\epsilon}_{t-q} \quad (1)$$

where $\vec{Z}_t, \vec{Z}_{t-k}, \dots, \vec{Z}_{t-p}$ are the vectors of values whose elements are random variables relating process values at time t, t-1, t-2, ..., t-p, $\vec{\epsilon}_t, \vec{\epsilon}_{t-1}, \dots, \vec{\epsilon}_{t-q}$ are independent vector random errors distributed with $N(0, G)$ at time t, t-1, ..., t-q respectively. $C_i, i=1, 2, \dots, p$ and $E_j, j=1, 2, \dots, q$ are coefficient matrices.

Two methods of estimating the parameters C_i , E_j , and G are available for the general multivariate model:

1. Method of Moments (O'Connell, 1974). This method is to use lag-zero, lag-one, lag-two cross-correlation matrices to derive coefficient matrices C_i , E_j , and covariance matrix G . It does not require any assumption as to the distribution from which the data is derived. On the other hand, it does not possess the desirable property of asymptotic efficiency (i.e., achieve the Cramer-Roe bound). Besides, it is impossible to correct for bias as is done by O'Connell (1974) with the method of moment in univariate series. Also, the estimation is restricted to the $p=1$, $q=1$ case.

2. Maximum Likelihood method (Box and Jenkins, 1970). This method requires the assumption of the distribution. The estimates are known to be asymptotically efficient and are considered better than those of O'Connell especially when moving average parameters are present (Ledolter, 1978). The method is virtually impossible to use in the general multivariate case due to a large number of parameters which require numerical optimization procedures for estimation.

The general multivariate model (1) possesses the characteristic that univariate (or subset) processes are not necessarily of the same order in both moving average and autoregressive components. Ledolter (1978) has shown that individual series from a multivariate AR process follow a univariate autoregressive model, but of higher order and with correlated residuals. The same can be shown for the moving average components. This results in the logical problem in which a univariate component of the model is not of the same order individually as in the multivariate

model. This causes a difficulty because the most natural method of understanding multivariate phenomenon is to first study the univariate cases and then generalize. For the unrestricted ARMA model, generalization is not necessarily possible.

Due to the weakness mentioned above, a restricted ARMA model is studied in this thesis. The model is defined to be one in which each univariate component of the multivariate model is of the same order in p and q as it is in the multivariate model or in any subset of the multivariate model which contains the component. Thus, we can follow the nature process to build models, starting with simpler univariate description and generalizing to the multivariate, based on information secured from the univariate study. The model is called a homogeneous model (HM).

The purpose of this project is two fold. First, to characterize the HM in terms of its parameteric and covariance structure. Second, to develop estimation procedures based on ML method and O'Connell's method for the univariate model. The model will be evaluated using the Monte Carlo technique only by ML method due to O'Connell's univariate estimation is difficult to apply on Monte-Carlo procedure. Those above will be accomplished in the following manner.

1. Generate homogeneous data and estimate the lag 0,1,2 correlation matrices using ML method to estimate the model parameters. Then compare the computed statistics with the same correlation matrices derived from the true model parameters. The purpose is to evaluate homogeneous model from homogeneous data by ML method.

2. Generate nonhomogeneous data and estimate the correlation matrices by ML method and compare these estimation values with the true values as in 1. The purpose is to evaluate how well the model works with data from a nonhomogeneous system.

In Chapter II, the homogeneous model is characterized from both theoretical and physical respects. The method of ML estimation for univariate model is described in Chapter III. O'Connell's univariate estimation method is described in Chapter IV. In Chapter V, the lag 0,1,2 correlation matrices are estimated by ML method from both homogeneous and nonhomogeneous data and compared with the corresponding statistics computed using the model parameters used to generate the data. O'Connell's general multivariate method of moment is tried here to evaluate ML method. The conclusions of this study are also summarized in this chapter.

CHAPTER II
CHARACTERIZATION OF MULTIVARIATE
HOMOGENEOUS ARMA MODEL

The homogeneous ARMA model as defined in Chapter I possesses the property of preserving the ARMA order of every subset of components of the system. In this chapter, the homogeneous model is characterized in both mathematical and physical respects.

Mathematical characterization

The following defines the homogeneous model in mathematical terms.

Theorem--The ARMA(p, q) model in Equation 1 is homogeneous if, and only if, the matrices $C_i, i=1, 2, \dots, p$ and $E_j, j=1, 2, \dots, q$ are diagonal.

Proof--If matrices C_i and E_j are diagonal, then as it is, every vector subset of the original variable has the model formed by taking the corresponding subset of the rows of the C_i and E_j matrices. Since these are diagonal, the subset matrix of the rows can be redefined to be diagonal and the $\vec{\epsilon}_t$ vector shortened to include only those elements necessary.

Conversely, if matrices C_i and E_j are not diagonal, suppose that one (or more) of the C_i and E_j have one or more off-diagonal elements that are nonzero, then it follows directly from the proof of the inconsistency referred to by Ledolter (1978) that there exists a marginal univariate model with autoregressive order greater than p .

For example, suppose there is one non-zero off-diagonal element of E_1 in the ARMA(1,1) model in the following equation:

$$\vec{Z}_t = C_1 \vec{Z}_{t-1} + \vec{\varepsilon}_t - E_1 \vec{\varepsilon}_{t-1} \quad (2)$$

Suppose it is the (1,2) element of E_1 , let b_1 represent the first row vector of E_1 . Then Equation 3 represents the first element of \vec{Z}_t , i.e., $z_{t,1}$

$$z_{t,1} = c_1 z_{t-1,1} + \varepsilon_{t,1} - e_1 \varepsilon_{t-1,1} - e_1' \varepsilon_{t-1,2} \quad (3)$$

Thus, the vector b_1 results in the inclusion of the second element of $\vec{\varepsilon}_{t-1} (e_1' \varepsilon_{t-1,2})$ into the model for $z_{t,1}$. Since this element cannot be incorporated in the "error" term at time t (i.e., $\varepsilon_{t,1}$), the error structure does not have the univariate moving average form.

It has been shown that no off-diagonal elements of $C_i, i=1,2,\dots,p$ nor $E_j, j=1,2,\dots,q$ can be nonzero or the ARMA(p,q) model will not preserve at least one univariate subset. Thus, the theorem is proven.

Maximum likelihood estimation for the nonhomogeneous model involves $n^2(p+q+1)$ parameters where n is the dimension of \vec{Z} in the model. Simultaneous numerical optimization is required which involves $n^2(p+q)$ parameters. Thus, for example (2), a system with five variables requires numerical optimization on 50 parameters. By contrast, the homogeneous model involves $n(p+q)$ parameters. For the five variable case, it is 10 parameters only. Numerical optimization on 50 variables seems questionable, while on 10 it is reasonable.

Physical characterization

It is important to investigate on intuitive grounds the effect of the restrictions as compared with the general ARMA models. The most notable effect due to the diagonal nature of the C_i and E_j matrices is that the historical contribution to the present value of a given

variable of the random vector \vec{Z} is limited to the historic value of that same variable. This does not at all imply independence because the present "error" (i.e., $\vec{\epsilon}_t$) can be a correlated random vector. As time progresses, e.g., as time t becomes $t+1$, the values in $\vec{\epsilon}_t$, as they contribute to the then present value of \vec{Z}_t , contribute to the new present value only as modified by a constant specific to each element in $\vec{\epsilon}_t$. For example, suppose the homogeneous model were used to model yearly flow volumes of several streams in a large basin. The restrictions imply that for a given stream, the present flow volume $Z_{i,t}$ can be molded by the present and past "errors" (ϵ_t and ϵ_{t-1}) and past streamflow in that stream drainage area ($Z_{i,t-1}$) and no other areas. The "errors" may be correlated, but only what actually happened in the past in the given area is what influences streamflow in that area. This implies that if, for example, one stream in the system receives considerable recharge from groundwater originating in the area of another stream in the system, the homogeneous model would not be applicable.

In the next chapter, estimation procedure by Maximum Likelihood is described and discussed.

CHAPTER III

MAXIMUM LIKELIHOOD ESTIMATION METHOD

Maximum Likelihood Estimation and O'Connell's method, i.e., bias-correct method of moments, are the two approaches used in this thesis to estimate parameters. In this chapter, the estimation procedure by the maximum likelihood method is described.

With the definition of Multivariate Homogeneous ARMA Model, the coefficient matrices C_i and E_j in Equation 1 are diagonal, and it possesses the property that the order of individual series contained within a multivariate model are of the same order in both p and q . Therefore, in this paper, univariate estimation by MLE is described here to estimate parameters $\phi_{i1}, \phi_{i2}, \dots, \phi_{im}$ and $\phi_{j1}, \phi_{j2}, \dots, \phi_{jm}$, i.e.,

$$C_i = \begin{bmatrix} \phi_{i1} & & & \\ & \phi_{i2} & & \\ & & \ddots & \\ 0 & \dots & & \phi_{im} \end{bmatrix} \quad i = 1, 2, \dots, p$$

$$E_j = \begin{bmatrix} \theta_{j1} & & & \\ & \theta_{j2} & & \\ & & \ddots & \\ 0 & \dots & & \theta_{jm} \end{bmatrix} \quad j = 1, 2, \dots, p$$

for the maximum likelihood method. Standard multivariate estimation by conditional MLE (Ledolter, 1978) is used to estimate the variance-covariance matrix G .

Maximum likelihood estimation

The maximum likelihood method of estimation is reviewed in this section. Let $\vec{Z}' = (z_1, z_2, \dots, z_n)$ be a sample of n independent observations from a population with density $f(z, \vec{R})$ where $\vec{R}' = (r_1, r_2, \dots, r_k)$ is a vector of parameters. The likelihood function is defined

$$L(\vec{Z}, \vec{R}) = \prod_{i=1}^n f(z_i, \vec{R})$$

Note that this is the joint density of the observations. It seems reasonable that the likelihood (i.e., probability density) of a sample using the true density should be high relative to that computed from any other density. Estimation by maximum likelihood chooses parameter values which maximize the likelihood function. When possible, the simplest method is to solve the set of equations constructed from the derivation of $l(\vec{Z}, \vec{R}) = \log(L(\vec{Z}, \vec{R}))$ with respect to the parameter equal to zero.

$$\frac{\partial l(\vec{Z}, \vec{R})}{\partial r_i} = 0 \quad i = 1, 2, \dots, k$$

Estimation of the parameters of the ARMA(p,q) is based upon this principle. The problem is more complex, however, and has given rise to two different methods originated by Box and Jenkins (1970). These methods are discussed in the following sections.

Conditional maximum likelihood estimation

Suppose that n original observations form a time series which we denote by z_1, z_2, \dots, z_n , assume that this series is generated by Equation

1. The mean μ can be estimated as $\bar{z} = \frac{1}{n} \sum_{t=1}^n z_t$, and a new sequence of

observations defined as $w_t = z_t - \bar{z}$, $E(w_t) = 0$ can be written as

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (4)$$

The term ϵ_t may be defined as

$$\epsilon_t = w_t - \phi_1 w_{t-1} - \phi_2 w_{t-2} - \dots - \phi_p w_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (5)$$

The w 's cannot be substituted immediately into Equation 5 to calculate the ϵ 's because of the difficulty in initiating the model in Equation 5. However, suppose that the p values of W^* of w 's and the q values of ϵ^* of the ϵ 's prior to the commencement of the w series were given. Then the values $\epsilon_i, i=1,2,\dots,n$ could be calculated. Assuming ϵ_i are normally distributed, the joint distribution (i.e., the likelihood function) of the $\vec{\epsilon}_t$ is

$$L(\epsilon_1, \epsilon_2, \dots, \epsilon_n) \approx \sigma_\epsilon^{-n} \exp(-(\sum \epsilon_t^2 / 2\sigma_\epsilon^2))$$

The log likelihood associated with the parameter values $(\vec{\Phi}, \vec{\Theta}, \sigma_\epsilon)$, conditional on the choice of w^* , ϵ^* (starting value) would then be

$$l^*(\vec{\Phi}, \vec{\Theta}, \sigma_\epsilon | w^*, \epsilon^*) = -n \log \sigma_\epsilon - \frac{S^*(\vec{\Phi}, \vec{\Theta})}{2\sigma_\epsilon^2} \quad (6)$$

where $\vec{\Phi}$ and $\vec{\Theta}$ denote the sets of parameters $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ and

$$\begin{aligned} S^*(\vec{\Phi}, \vec{\Theta}) &= \sum \epsilon_t^2(\vec{\Phi}, \vec{\Theta} | w^*, \epsilon^*) \\ &= \sum (w_t - \phi_1 w_{t-1} - \phi_2 w_{t-2} - \dots - \phi_p w_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q})^2 \end{aligned}$$

The conditional likelihood $L(\vec{\Phi}, \vec{\Theta}, \sigma_\epsilon^2)$ involves the data only through the conditional sum of squares function. It follows from the normal

assumption, that values of the elements of $\vec{\Phi}$ and $\vec{\Theta}$ which minimize $S^*(\vec{\Phi}, \vec{\Theta})$ maximizing $L(\vec{\Phi}, \vec{\Theta}, \sigma_\varepsilon^2)$ and the resulting least squares estimated represent ML estimates.

Unconditional maximum likelihood estimation

The unconditional log-likelihood function for a sequence of n observation assumed to have been generated by Equation 4 is given as

$$l(\vec{\Phi}, \vec{\Theta}, \sigma_\varepsilon^2) = f(\vec{\Phi}, \vec{\Theta}) - n \log \sigma_\varepsilon - \frac{S(\vec{\Phi}, \vec{\Theta})}{2\sigma_\varepsilon^2} \quad (7)$$

where $f(\vec{\Phi}, \vec{\Theta})$ is a function of Φ and Θ . The unconditional sum of squares function is given as

$$S(\vec{\Phi}, \vec{\Theta}) = \sum_{t=-\infty}^n \{\varepsilon_t | \vec{\Phi}, \vec{\Theta}, \vec{W}\}^2 \quad (8)$$

where $\{\varepsilon_t | \vec{\Phi}, \vec{\Theta}, \vec{W}\} = E\{\varepsilon_t | \vec{\Phi}, \vec{\Theta}, \vec{W}\}$, $E\{\varepsilon_t | \vec{\Phi}, \vec{\Theta}, \vec{W}\}$ denotes the expectation of ε_t conditional on $\vec{\Phi}, \vec{\Theta}, \vec{W}$ which is different from Conditional Likelihood estimation as conditional on starting value W^*, ε^* . As $f(\vec{\Phi}, \vec{\Theta})$ is usually unimportant for moderate to large n , (Box and Jenkins, 1970) contours of $S(\vec{\Phi}, \vec{\Theta})$ closely approximate contours of log-likelihood. Hence, least squares estimates obtained through minimizing $S(\vec{\Phi}, \vec{\Theta})$ in Equation 7 will usually provide close approximation to ML estimates.

Ideally, the unconditional likelihood function should be used for parameter estimation (Box and Jenkins, 1970), but suitable choices of W^*, ε^* allow a sufficient approximation to the unconditional likelihood function for moderate to large n by using the conditional likelihood function. One choice for the elements of ε^* and W^* would be the unconditional expectation of ε_t and w_t which are zero. However, if the

values of some of the autoregressive parameters lie near boundaries, then this approximation may not be sufficient. A more reliable procedure is to calculate the values of ε_t from Equation 5 for ε_{p+1} $t=1,2,\dots$ setting previous ε_t values to zero. Hence, the sum of squares S^* will then be derived from $(n-p)$ values of ε_t , but the slight loss of information would be unimportant for long series (Box and Jenkins, 1970). However, for short series, the best approach is to work with the unconditional log-likelihood function.

General procedure for calculating the unconditional sum of squares

While the conditional expectation ε_t is linear in the elements of the parameters set $\vec{\Phi}$, it may be shown to be nonlinear in the elements of the set $\vec{\Theta}$. Consequently, techniques which rely on $S(\vec{\Phi}, \vec{\Theta})$ being quadratic in the parameters, such as linear least squares, are not strictly applicable. Box and Jenkins (1970) suggest how ε_t may be suitably linearized and how linear least squares techniques may then be applied iteratively to obtain ML estimates, provided reasonable initial guess at the parameter values are available. However, more general optimization techniques for finding the greatest or least value of a function without calculating derivatives are now widely available, and may be applied to minimize the function $S(\vec{\Phi}, \vec{\Theta})$.

In evaluating Equation 8, conditional expectations are taken in Equation 5 to yield a value of $\{\varepsilon_t | \vec{\Phi}, \vec{\Theta}, \vec{w}\}$. As the recurrence relationship starts with $t=1$ and proceeds forward, values $w_{-j}, j=0,1,\dots$ are required to start off the forward recurrence relationship. In order to provide those values, Box and Jenkins (1970) define ARMA(p,q) process using a forward shift operator

$$Fw_t = w_{t+1}$$

and define the value of the w_t at time t in terms of a set of random shocks $\delta_t, \delta_{t+1}, \delta_{t+2}, \dots$

$$\Phi(F)w_t = \Theta(F)\delta_t$$

which may be written in recurrent form as

$$\delta_t = w_t - \phi_1 w_{t+1} - \phi_2 w_{t+2} - \dots - \phi_p w_{t+p} + \theta_1 \delta_{t+1} + \theta_2 \delta_{t+2} + \dots + \theta_q \delta_{t+q} \quad (9)$$

As $w_t, t=1, 2, \dots, n$ are distributed independently of $\delta_t, t=0, -1, -2, \dots$ then $\delta_0 = \delta_{-1} = \delta_{-2} = \dots = 0$, set $\delta_n = 0$.

Rearranging Equation 9 as

$$w_t = \delta_t + \phi_1 w_{t+1} + \phi_2 w_{t+2} + \dots + \phi_p w_{t+p} - \theta_1 \delta_{t+1} - \theta_2 \delta_{t+2} - \dots - \theta_q \delta_{t+q}$$

the values $w, w_{-1}, w_{-2}, \dots, w_{-m}$ may be calculated until w_{-m} has become sufficiently small as $|\phi_i| < 1$. (Refer examples of Box and Jenkins, 1970, p. 215-219).

Using Equation 5, values $\varepsilon_{-m}, \varepsilon_{-m+1}, \dots, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_n$ are calculated assuming that w_{-m} is effectively zero, which means that $\varepsilon_{-j} \approx 0$, for $j > m-1$. The unconditional sum of squares $S(\vec{\Phi}, \vec{\Theta})$ is obtained through summing the squares of all the calculated ε_t values.

Estimation of variance-covariance matrix

Regarding the estimation of variance-covariance matrix G , the method of multivariate conditional maximum likelihood estimated suggested by Ledolter (1978) is a generalization of the iterative estimation procedure originated by Box and Jenkins (1970) for the univariate case.

Let $\vec{\beta}$ be a column vector of unknown parameters $\vec{\phi}_1, \vec{\phi}_2, \dots, \vec{\phi}_p, \vec{\theta}_1, \vec{\theta}_2, \dots, \vec{\theta}_q$.

$$\vec{\phi}_i = \begin{bmatrix} \phi_{i1} \\ \vdots \\ \phi_{im} \end{bmatrix}, \quad \theta_j = \begin{bmatrix} \theta_{j1} \\ \vdots \\ \theta_{jm} \end{bmatrix} \quad \begin{array}{l} i = 1, 2, \dots, p \\ j = 1, 2, \dots, q \end{array}$$

and G is the variance-covariance matrix, then the multivariate maximum likelihood functions can be developed from Equation 5 as follows:

$$L(\vec{\beta}, G | \vec{Z}_1, \vec{Z}_2, \dots, \vec{Z}_n) = |G|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \vec{\epsilon}'(t) G^{-1} \vec{\epsilon}(t) \right\}$$

The log-likelihood function is then

$$l(\vec{\beta}, G | \vec{Z}_1, \vec{Z}_2, \dots, \vec{Z}_n) = -\frac{n}{2} \left\{ \log |G| + \frac{1}{n} \sum_{t=1}^n \vec{\epsilon}'(t) G^{-1} \vec{\epsilon}(t) \right\}$$

The conditional estimate of G which was derived by Ledolter (1974) is given by

$$G = \frac{1}{n} \sum_{t=1}^n \vec{\epsilon}'(t) \vec{\epsilon}(t) \quad (10)$$

In the next chapter, O'Connell's estimation method is introduced.

CHAPTER IV

O'CONNELL'S ESTIMATION METHOD

Besides ML method, another estimation method for the multivariate homogeneous model uses O'Connell's (1974) univariate estimation by the method of moments with a bias correction. In this chapter, this method is reviewed and discussed.

The O'Connell's multivariate model as shown below is parameterized differently from the general multivariate model shown in Equation 1. O'Connell's model is

$$\vec{X}_t = A\vec{X}_{t-1} + B\vec{\epsilon}_t^0 - C\vec{\epsilon}_{t-1}^0 \quad (11)$$

The differences are, first, the model is limited to $p=1, q=1$, i.e., ARMA (1,1) due to difficulties of estimating higher order terms. Second, elements in the model are represented differently, i.e., $\epsilon_t^0, \epsilon_{t-1}^0$ are independent vectors random errors distributed with $N(0, I)$ at time $t, t-1$. Also \vec{X}_t, \vec{X}_{t-1} are standardized vectors in which $x_t = (x_t - u)/\sigma$, $x_{t-1} = (x_{t-1} - u)/\sigma$, u is the mean of the process, σ is the series variance, $\vec{A}, \vec{B}, \vec{C}$ are coefficient matrices with respect to $\vec{X}_{t-1}, \vec{\epsilon}_t^0, \vec{\epsilon}_{t-1}^0$. However, this model can be adjusted to correspond to general ARMA(1,1), model (2) in the following way:

$$\text{Let } D\vec{X}_t = \vec{Z}_t \quad D = \text{diag}(\sigma_1 \sigma_2)$$

Multiplying Equation 11 by D , gives

$$D\vec{X}_t = DA\vec{X}_{t-1} + DB\vec{\epsilon}_t^0 - DC\vec{\epsilon}_{t-1}^0$$

equals

$$\vec{Z}_t = A\vec{Z}_{t-1} + \vec{\epsilon}_t - CB^{-1}\vec{\epsilon}_{t-1}$$

so, model

<u>Equation 1</u>	=	<u>Equation 2</u>
A	=	C
CB^{-1}	=	E
$(DB)'(DB)$	=	G

which shows to estimate parameter A,B,C is equivalent to estimate parameter C,E in Equation 2. Therefore, in this paper, O'Connell's univariate estimation may be used to estimate parameters $\phi_1, \phi_2, \dots, \phi_m$ and $\theta_1, \theta_2, \dots, \theta_m$ for C and E, i.e.

$$C = \begin{bmatrix} \phi_1 & \dots & 0 \\ & \phi_2 & \\ & & \ddots \\ 0 & \dots & \phi_m \end{bmatrix}, \quad E = \begin{bmatrix} \theta_1 & \dots & 0 \\ & \theta_2 & \\ & & \ddots \\ 0 & \dots & \theta_m \end{bmatrix}$$

and also standard multivariate estimation by conditional MLE (Ledolter, 1978) as used in Chapter III is suggested here to estimate the covariance matrix G.

The most significant difference between O'Connell's univariate estimation and the MLE is that it incorporates consideration of a measure of long-term persistence h , called the Hurst Coefficient.

Persistence is the tendency for a high observation (e.g., stream-flow) to be followed by high observations and a low outcome to be followed by low outcomes. Hurst (1951, 1956) made the remarkable empirical discovery that a host of geophysical time series obeyed one universal, probabilistic law, specified by one parameter, $0 < h < 1$, which governed the duration and intensity of periods of above and below average outcomes.

Two methods are used to estimate Hurst Coefficient h :

Method I

Suppose x_1, x_2, \dots, x_n denote a sequence of data over n years. Let \bar{x} denote the sample mean and define

$$D_n = \sum_{i=1}^n x_i$$

Let k be the first k years.

$$D_k^* = \sum_{i=1}^k x_i - \frac{k}{n} \sum_{i=1}^n x_i$$

represent the excess or deficiency relative to the amount removed up to the k^{th} year. Defining

$$M_k = \max D_k^* \quad 1 \leq k \leq n$$

$$m_k = \min D_k^* \quad 1 \leq k \leq n$$

as the largest excess and greatest deficiency respectively over the steady outcome during the n years. The quantity

$$R = M_k - m_k$$

is known as the range of cumulative departures from the sample mean.

Hurst found R to vary with n as

$$R/S \approx n^h$$

where S denotes the sample standard deviation of the time series of length n , h is a constant, R/S is referred to as the rescaled range. The coefficient h was estimated through the following relationship.

$$\frac{R}{S} = \left(\frac{n}{2}\right)^K$$

where K denotes the resulting estimate of the population coefficient h.

Thus, K was defined for each time series as

$$K = \frac{\log R - \log S}{\log n - \log 2}$$

Hurst (1951, 1956) collected 900 annual time series comprising stream-flow and precipitation ... etc. records, and found values of K which ranged from .46 to .96 with a mean of .729 and a standard deviation of .092 over all phenomena.

Method II

The second method was proposed by Mandelbrot and Wallis (1969b) which provides an estimate denoted by H. This method consists of the following three steps:

1) A record of length n is divided into N subsamples of length, n_s , where $3 < n_s < n$. The subsamples may or may not overlap, and the selection of the subsample length n_s is made such that a uniform spacing of n_s is achieved on a logarithmic scale.

2) The rescaled range R/S is computed for each subsample of length n_s .

3) A least square lines is fitted to the mean R/S for each subsample size in the range $n_0 < n_s < n$, where n is chosen on the grounds of an initial non-linear transient in the plot of $\log(R/S)$ against $\log n$.

The slope of the fitted least square line yields the estimate H.

This diagram is referred to as the "pox-diagram."

Bias correction

However, it was found by Matales and Wallis (1970) that both K and H are biased estimators, with K displaying greater bias than H ; but K was found to have smaller variance than H . The bias in both estimators decreases slowly with n , the sample size.

The problem of bias correction was circumvented by O'Connell (1974) through defining $E(K)_n$ or $E(H)_n$, the expected value of K or H in sample of size n for a process. Invariably, Monte-Carlo simulation is necessarily used to define $E(H)_n$ or $E(K)_n$ for a process.

In defining the small sample properties of estimates of h for a process, a choice must be made between H and K . Since H suffers from a major deficiency in that no universal rule exists for defining the set of sub-series to be used, or for estimating the slope of the "pox-diagram," and K does not suffer from such a deficiency and is also quicker to compute. Accordingly, K was used by O'Connell (1974) in the simulation experiments to define $E(K)_n$ for ARMA(1,1) process.

Agreement with Hurst Law

To examine the ability of ARMA(1,1) process to model Hurst's law accurately for values of n comparable with the longest geophysical records available, a simulation approach was adopted and a detailed account of the simulation experiments conducted for this purpose has been presented by O'Connell (1971). Over 250 "pox-diagrams" corresponding to a number of combinations of ϕ and θ were constructed, and good overall agreement with Hurst's law was observed up to moderate to large value of n .

The simulation experiments for defining $E(K)_n$ were extended to large values of n . Sample of size $n = 25, 50, 100, 250, 500, 1000, 2500, 5000, 10000$ were generated and estimates of the lag-one autocorrelation ρ_1 and K were made for each sequence. The estimation of K has been described above. The lag-one autocorrelation is estimated by

$$\hat{\rho}_1 = \frac{\left\{ \sum_{i=1}^{n-1} x_i - \frac{1}{n-1} \sum_{i=1}^{n-1} x_i \right\} \left\{ \sum_{i=1}^{n-1} x_{i+1} - \frac{1}{n-1} \sum_{i=1}^{n-1} x_{i+1} \right\}}{\left\{ \sum_{i=1}^{n-1} \left(x_i - \frac{1}{n-1} \sum_{i=1}^{n-1} x_i \right)^2 \right\} \left\{ \sum_{i=1}^{n-1} \left(x_{i+1} - \frac{1}{n-1} \sum_{i=1}^{n-1} x_{i+1} \right)^2 \right\}}^{1/2}$$

and for the ARMA(1,1) model it is equal to

$$\rho_1 = \frac{(\phi - \theta)(1 - \phi\theta)}{(1 - \theta^2 - 2\phi\theta)} \quad (13)$$

Estimates of $E(K)_n$ and $E(\hat{\rho})_n$, denoted by $\tilde{E}(K)_n, \tilde{E}(\hat{\rho})_n$ were then derived by averaging the respective statistics over the total number of realizations.

For selection of values of ϕ and θ , tables of $E(K)_n$ and $E(\rho_1)$ have been abstracted for sample size $n = 25, 50, 100$. The value of $E(K)_n$ and $E(\rho_1)_n$, each derived from 10,000 samples of size n are presented by O'Connell (1979) in Tables 4.1-4.6 which are reproduced in Appendix C. For sample size 25, the range of K values which can be modelled for all the selected values of ϕ and θ is approximately .65-.80 while for sample size 50 and 100, the corresponding range are .65-.85 and .65-.87 respectively.

The procedure to estimate parameter ϕ_1 and $\theta_i, i=1, 2, \dots, m$ is as follows:

1. Derive estimates ρ_1, K from a historic sequence of length n .
2. From O'Connell's Tables 4.1-4.6, identify values of ϕ and θ such that

$$\tilde{E}(K)_n \approx \rho$$

and

$$\rho_1 \approx \hat{\rho}_1$$

where ρ_1 is defined by Equation 13.

Despite the advantage of O'Connell's univariate estimation, that is the bias correction to ensure the presentation of the Hurst Coefficient, it is difficult to apply Monte-Carlo simulation to O'Connell's procedure. This is because no direct relationship can be found between K , ρ , ϕ , and θ . Therefore, in the next chapter, a two-dimensional homogeneous model is to be evaluated only by method of MLE. Monte-Carlo methods will be used to measure the mean squared deviation of the sample lag 0,1 and 2 correlation matrices from the population values. O'Connell's multivariate estimation method of moments was attempted in order to compare it with the estimation by Maximum Likelihood, but no satisfactory result was obtained due to difficulties encountered in estimation. Limited results using O'Connell's method are reported in the next chapter.

CHAPTER V

EVALUATION OF HOMOGENEOUS MODEL

Introduction

ML method of estimation described in Chapter III is used in this chapter to estimate parameters in a Monte-Carlo experiment. The homogeneous model is evaluated by generating different sample sequence using a known model, estimating the parameters by ML method, then comparing the results in the following ways:

- (a) Compare the estimated model parameters with the respective population parameters.
- (b) Compare the sample lag 0,1 and 2 correlation matrices as computed from estimated model parameters with the same parameters computed from the population model parameters.

From these comparisons, how well the homogeneous model works on homogeneous data or nonhomogeneous data can be observed. The correlation matrices constitute a measure of multivariate model performance, also, the mean square error shows the precision of estimation.

Following is the description of the program in this thesis. The Monte-Carlo simulation is used to evaluate homogeneous model from different types of data. For simplicity, ARMA(1,1) model with a two-dimensional variable is simulated.

Let the model ARMA(1,1) adjusted from Equation 2 into the form of Equation 4 as shown below:

$$\vec{W}_t = C\vec{W}_{t-1} + \vec{\epsilon}_t - E\vec{\epsilon}_{t-1} \quad (14)$$

Let the $\phi_{11} = .9$, $\phi_{22} = .8$ and $\theta_{11} = .4$, $\theta_{22} = .2$, i.e., in Equation 14

$$(a) \quad C = \begin{bmatrix} .9 & 0 \\ 0 & .8 \end{bmatrix} \quad E = \begin{bmatrix} .4 & 0 \\ 0 & .2 \end{bmatrix}$$

be the parameters used to generate homogeneous data. Three other sets of coefficient matrices C,E are used to generate data reflecting different degrees of unhomogeneity as shown below.

$$(b) \quad C = \begin{bmatrix} .9 & 1 \\ .1 & .8 \end{bmatrix} \quad E = \begin{bmatrix} .4 & .1 \\ .1 & .2 \end{bmatrix}$$

$$(c) \quad C = \begin{bmatrix} .9 & .4 \\ .4 & .9 \end{bmatrix} \quad E = \begin{bmatrix} .4 & .1 \\ .1 & .2 \end{bmatrix}$$

$$(d) \quad C = \begin{bmatrix} .9 & .7 \\ .7 & .8 \end{bmatrix} \quad E = \begin{bmatrix} .4 & .1 \\ .1 & .2 \end{bmatrix}$$

Let model variance-covariance matrix be defined as

$$G = \begin{bmatrix} \sigma_1^2 & .5\sigma_1\sigma_2 \\ .5\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

From Box and Jenkins (1970), it follows that

$$\sigma_1^2 = (1 - \phi_{ii})^2 / (1 + \theta_{ii}^2 - 2 \phi_{ii} \theta_{ii}) \quad i=1,2$$

so that

$$G = \begin{bmatrix} .432 & .2324 \\ .2324 & .5 \end{bmatrix}$$

To construct synthetic data from Equation 14, let $SS' = G$, $S = UV^{1/2}$, where V is the matrix with each diagonal element representing an individual characteristic root, $V^{1/2}$ is obtained from the square root of diagonal elements, and U is the matrix of characteristic vector corresponding to each characteristic root of matrix G . ϵ is the generating random number distributed with $N(0, I)$, so $S\vec{\epsilon} = \vec{\epsilon}$ is then distributed with $N(0, G)$, since

$$E\{(\vec{S\epsilon})(\vec{S\epsilon})'\} = E(\vec{S\epsilon\epsilon'}S') = SE(\vec{\epsilon\epsilon'})S' = SIS' = G$$

Using the parameters C, E defined above and the random error sequence $\epsilon_t, \epsilon_{t-1}, t=1, 2, \dots, 80$, 400 series each of length 80 is generated.

In order to prohibit the nonstationary sequence which may occur at the beginning of generation, the first 19 data are deleted from whole length, and replace the 1st data by the 20th, then start all over again. So the 80 data is actually generated beginning with 20th. For each series of data, ML method as described in Chapter III is used to estimate parameter C, E , and G . In the program, subroutine FTMXL (IMSL, 1979),

which is a program of univariate, conditional MLE to estimate C and E, is used.

Expected values of correlation matrices

Lag 0, 1 and 2 correlation matrices are to be computed directly from those parameters. It is shown below how to obtain the direct relationship between parameters and correlation matrices M_0, M_1, M_2 .

Let $\vec{w}_t, \vec{w}_{t-1}, \dots, \vec{w}_{t-n+1}$ be a series of n observations from model 14, then

$$\vec{w}_t = C\vec{w}_{t-1} + \vec{\epsilon}_t - E\vec{\epsilon}_{t-1}$$

$$\vec{w}_{t-1} = C\vec{w}_{t-2} + \vec{\epsilon}_{t-1} - E\vec{\epsilon}_{t-2}$$

... ..

$$\vec{w}_{t-n+1} = C\vec{w}_{t-n} + \vec{\epsilon}_{t-n+1} - E\vec{\epsilon}_{t-n}$$

Starting from the bottom, replace the \vec{w}_{t-n+1} in the right side of the next to last equation by its expression in the last equation. Then proceed upward, replacing the \vec{w}_{t-n+i} on the right by the previously derived expression. The final result for \vec{w}_t is

$$\vec{w}_t = C(C\vec{w}_{t-2} + \vec{\epsilon}_{t-1} - E\vec{\epsilon}_{t-2}) + \vec{\epsilon}_t - E\vec{\epsilon}_{t-1}$$

$$= C^2\vec{w}_{t-2} - CE\vec{\epsilon}_{t-2} + (C-E)\vec{\epsilon}_{t-1} + \vec{\epsilon}_t$$

$$= C^2(C\vec{w}_{t-3} - \vec{\epsilon}_{t-2} - E\vec{\epsilon}_{t-3}) - CE\vec{\epsilon}_{t-2} + (C-E)\vec{\epsilon}_{t-1} + \vec{\epsilon}_t$$

.....

$$\begin{aligned}
&= C^n \vec{w}_{t-n} + C^n (C-E) \vec{w}_{t-n+1} + \dots + C^2 (C-E) \vec{\varepsilon}_{t-3} + C (C-E) \vec{\varepsilon}_{t-2} \\
&+ (C-E) \vec{\varepsilon}_{t-1} + \vec{\varepsilon}_t
\end{aligned}$$

As n goes to infinity, the result is

$$\vec{w}_t = \vec{\varepsilon}_t + (C-E) \vec{\varepsilon}_{t-1} + C(C-E) \vec{\varepsilon}_{t-2} + C^2(C-E) \vec{\varepsilon}_{t-3} + \dots$$

From $E(\vec{w}_t \vec{w}_t')$, the lag 0 variance-covariance matrix V_0 is

$$\begin{aligned}
V_0 = E(\vec{w}_t \vec{w}_t') &= C^{n-1} E G (C^{n-1})' + C^{n-2} (C-E) G (C-E)' C^{n-2} + \dots + \\
&D(C-E) G (C-E)' C + (C-E) G (C-E)' + G
\end{aligned} \quad (15)$$

The lag 1 variance-covariance matrix V_1 is

$$\begin{aligned}
V_1 = E(\vec{w}_t \vec{w}_{t-1}') &= C^{n-1} E G (C-E)' (C^{n-2})' + C^{n-2} (C-E) G (C-E)' (C^{n-3})' + \dots + \\
&C(C-E) G (C-E)' + (C-E) G
\end{aligned} \quad (16)$$

The lag 2 variance-covariance matrix V_2 is

$$\begin{aligned}
V_2 = E(\vec{w}_t \vec{w}_{t-2}') &= C^{n-1} G (C-E)' (C^{n-3})' + C^{n-2} (C-E) G (C-E)' (C^{n-4})' + \dots + \\
&C(C-E) G
\end{aligned} \quad (17)$$

Since lag 0,1,2 variance-covariance matrices are defined as following forms:

$$V_0 = E(\vec{w}_t \vec{w}_t') = \begin{bmatrix} \sigma_1^2 & \dots & & & \\ & \ddots & \sigma_2^2 & & \\ & & \ddots & \ddots & \\ & & & \rho_{ij} \sigma_i \sigma_j & \\ & & & & \ddots \\ \rho_{ji} \sigma_j \sigma_i & \dots & & & & \sigma_m^2 \end{bmatrix}$$

In 2x2 case

$$V_0 = \begin{bmatrix} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 \\ \rho_{21} \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$\text{where } E(w_i w_j) = \sigma_i^2 \quad i=j$$

$$E(w_i w_j) = \rho_{ij} \sigma_i \sigma_j \quad i \neq j$$

$$M_1 = \begin{bmatrix} \rho_1(1) & & & \\ & \ddots & & \\ & & \rho_{ij}(1) & \\ \rho_{ji}(1) & & & \\ & & & \ddots \\ & & & & \rho_m(1) \end{bmatrix} \quad M_1 = \begin{bmatrix} \rho_1(1) & \rho_{12}(1) \\ \rho_{21}(1) & \rho_2(1) \end{bmatrix}$$

$$M_2 = \begin{bmatrix} \rho_1(2) & & & \\ & \ddots & & \\ & & \rho_{ij}(2) & \\ \rho_{ji}(2) & & & \\ & & & \ddots \\ & & & & \rho_m(2) \end{bmatrix} \quad M_2 = \begin{bmatrix} \rho_1(2) & \rho_{12}(2) \\ \rho_{21}(2) & \rho_2(2) \end{bmatrix}$$

Results

Table 1 shows the model parameters and the lag 0,1,2 correlation matrices used to generate data. In addition to the bias of M_0, M_1, M_2 , a measure of the precision of estimation is needed. Therefore, the mean square error (MSE) is used here which is defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - Y)^2$$

where Y_i represents estimated value from each iteration, Y represents the true value. From MSE, the variation of estimated value to true value can be observed. In Table 2, all the estimated parameters and statistics are listed which includes coefficient matrices C, E and their MSE to initial parameter, variance-covariance matrix G with its MSE to original G , lag 0,1,2 correlation matrices and their MSE from true statistics.

As mentioned before, due to difficulties experienced with O'Connell's multivariate method of estimation, only limited results are included here for comparison. The difficulty encountered is that

Table 1. Population model parameters C, E, and G and their lag 0,1,2 correlation matrices M_0, M_1, M_2 .

	Homogeneous				Non-Homogeneous			
	(a)		(b)		(c)		(d)	
C	.9	0	.9	.1	.9	.4	.9	.7
	0	.8	.1	.8	.4	.8	.7	.8
E	.4	0	.4	.1	.4	.1	.4	.1
	0	.2	.1	.2	.1	.2	.1	.2
G	.432	.2324	.432	.2324	.432	.2324	.432	.2324
	.2324	.5	.2324	.5	.2324	.5	.2324	.5
M_0	1.0	.30212	1.0	.302952	1.0	.358	1.0	.50603
	.30212	1.0	.302952	1.0	.358	1.0	.50603	1.0
M_1	.72644	.1162	.7270547	.123294	.83564	.26145	.88754	.466734
	.1394	.6999	.146544	.6999	.2634	.77741	.45894	.8208
M_2	.653799	.1046	.6690098	.180994	.848457	.51978	1.0692	.9037
	.11155	.5599	.1899102	.572324	.5761	.736244	1.1052	1.04475

Table 2. Estimated parameters C, E, and G and their MSE to population parameters. Estimated lag 0,1,2 correlation matrices and their MSE to population correlations.

	(a)		(b)		(c)		(d)	
C	.8144323	0	.8794567	0	1.245452	0	1.535315	0
	0	.7366728	0	.8056161	0	1.245453	0	1.535315
MSE	.02436553	0	.00946404	.01	.1193371	.16	.4036257	.49
	0	.0218169	.01	.01483797	.16	.1984285	.49	.5406889
E	.3261681	0	.3692978	0	-.9268363	0	-.9265654	0
	0	.1584256	0	.253059	0	-.9268298	0	-.9265652
MSE	.03099633	0	.02135154	.01	1.760495	.01	1.759776	.01
	0	.0341059	.01	.0342053	.01	1.269745	.01	1.2697149
G	.4334362	.2248046	.4387916	.2182003	3.571578	2.667185	.115666E+11	.103944E+10
	.2248046	.5023816	.2182003	.488983	2.667185	2.453855	.103944E+10	.101338E+11
MSE	.00448586	.00287923	.004899293	.002979122	9.881922	5.939306	.235412E+21	.714219E+20
	.00287923	.00627296	.002979122	.006141037	5.939306	3.828263	.714219E+20	.154489E+21
M ₀	1.0	.3318697	1.0	.2733108	1.0	.00000178	1.0	-.58E-12
	.3318697	1.0	.2733108	1.0	.00000178	1.0	-.58E-12	1.0
MSE	0	.006950614	0	.008332	0	.12688	0	.25479
	.006950614	0	.008332	0	.12688	0	.25479	0
M ₁	.6375	.1233457	.7177833	.1052457	1.245452	.6764E-6	1.535315	-.2026E-12
	.1476681	.6485394	.1157202	.6790808	.6764E-6	1.245453	-.2026E-12	1.535315
MSE	.02618064	.00106575	.018265	.00137	.1679877	.054834	.419678	.2178406
	.00164085	.01334956	.0026028	.01476698	.069374	.219065	.21064	.5105325
M ₂	.5304899	.09954137	.6393536	.09086183	1.551151	.84242E-5	2.357193	-.3111E-12
	.1068059	.4857257	.09041459	.5557084	.8424E-5	1.551154	-.3111E-12	2.357194
MSE	.0405236	.0007394	.0271172	.00087388	.4938	.270097	1.658928	.8176725
	.00088768	.0241686	.01065	.02336236	.331813	.66388	1.221467	1.72214

an excessive number of series, there is no real solution to estimation equations. This problem is more pronounced as the degree of nonhomogeneity increases. Monte-Carlo results for estimation by O'Connell's method are shown in Table 3. The sample size is 20, as compared with 400 for the MLE table.

It is noted that the true lag-two correlations are greater than 1 for model (d) Table 1. These inadmissible values could be from sources:

- (1) The off-diagonal elements on C are too big as .7.
- (2) The n chosen in Equation 17 is 30 which may not be large enough to provide a good approximation. Since the objective of this study has been achieved without resolving this problem, it has been ignored.

The estimated lag 1,2 correlations exceed 1 on (c) and (d) type of data sets is due to the bad estimator C and E, which also shows the instability of homogeneous model.

Conclusions and recommendations

The problem of estimation inherent with both MLE and the method of moments for the general multivariate model makes the homogeneous model attractive. It appears to be the only model that is computationally feasible for multivariate, multilag systems. Multiple lags are required to reproduce series where lag-two (or higher) correlations exceed lag-one correlations because of long aquifer travel time or long carryover storage periods in large reservoirs.

It is noted that the homogeneous model is sensitive to departures from the homogeneous assumptions, i.e., when the data originates from

Table 3. Estimated parameters C, E, and G and their MSE to population parameters. Estimated lag 0,1,2 correlation matrices and their MSE to population correlations by O'Connell's method.

C	.7892882	.07732155	.8212594	.1465434
	-.04119488	.757	.07925378	.7385913
MSE	.03754279	.02758439	.01745764	.02053561
	.0286879	.02737192	.01472264	.03891678
E	.2994335	-.018552	.2720511	.05514694
	.05959	.145	.1429307	.154423
MSE	.02889364	.04421245	.031847	.039198
	.03345558	.0555102	.03121261	.06171842
G	.4043	.2107796	.4596424	.2807509
	.2107796	.5833343	.2807509	.6240483
MSE	.00594367	.0064046	.0071054	.010151
	.0064046	.09031583	.010151	.06642188
M_0	1.0	.2730639	1.0	.3207595
	.2730639	1.0	.3207595	1.0
MSE	0	.008625715	0	.0087083
	.008625715	0	.0087083	0
M_1	.676948	.1459023	.7320732	.1843177
	.0716889	.6658066	.1107831	.6611064
MSE	.02543985	.0056932	.010902	.008725
	.0067287	.03	.005645	.02725193
M_2	.5645263	.3648067	.6333536	.3293157
	.058154	.5235498	.1644624	.5324146
MSE	.04941342	.3411693	.0188313	.08155
	.01649328	.05499427	.015552	.043664

a nonhomogeneous source. Small departures seem to cause no serious problem, however, large departures are serious. This lack of robustness underscores the need for a test of homogeneity. If such a test could be developed, it would serve to screen out data sets which are not computable with the model.

REFERENCES

Box, G. E. P., and Jenkins, G. M., Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day Inc., 1970.

International Mathematical Statistical Library, Houston, Texas.

James, L. Douglas, Bowles, David S., James, W. Robert, and Canfield, Ronald V., Estimation of Water Surface Elevation Probabilities and Associated Damages for the Great Salt Lake, Utah Water Research Laboratory, College of Engineering, Utah State University, 1979.

Ledolter, J. "The Analysis of Multivariate Time Series Applied to Problem in Hydrology." J. Hydrol. 36 (1978): 327-352.

O'Connell, P. E., "Stochastic Modelling of Long-term Persistence in Streamflow Sequence," Civil Engineering Department, Imperial College, London, 1974-2.

APPENDICES

Appendix A

Computer Program in Maximum

Likelihood Method

Input Data

- | | |
|---------------|--|
| 1. NX | sample size (# of series) |
| 2. SEED | value used to generate random numbers |
| 3. MEAN(I) | overall mean of each variable |
| 4. SCALE(I) | the conversion factor to maintain all variables in same unit |
| 5. OI(I,J) | coefficient elements in autoregressive term |
| 6. PI(I,J) | coefficient elements in moving average term |
| 7. S(I,J) | square roots of variance-covariance matrix |
| 8. ASM(I,J) | true lag 0 correlation matrix |
| 9. ASM1(I,J) | true lag 1 correlation matrix |
| 10. ASM2(I,J) | true lag 2 correlation matrix |
| 11. HOV(I,J) | true variance-covariance matrix |

Output Data

- | | |
|----------|---|
| 1. M0 | lag 0 correlation matrix |
| 2. M1 | lag 1 correlation matrix |
| 3. M2 | lag 2 correlation matrix |
| 4. CMSE1 | mean square error of estimated M0 to true M0 |
| 5. CMSE1 | mean square error of estimated M1 to true M1 |
| 6. CMSE3 | mean square error of estimated M2 to true M2 |
| 7. EA | coefficient matrix in autoregressive term (C) |
| 8. EB | coefficient matrix in moving average term (E) |
| 9. MSE1 | mean square error of estimated C to true C |
| 10. MSE2 | mean square error of estimated E to true E |
| 11. AVCV | variance-covariance matrix G |
| 12. CVMS | mean square error of estimated G to true G |

T H E S I S
= = = = =

FILE 6(KIND=DISK)
FILE 16(KIND=PRINTER)

REAL MSE1,MSE2
INTEGER SEED
INTEGER G,W,H
REAL MD,M1,M2
REAL MLT
DIMENSION BT(100,2),MLT(2,2),WT(2,2),T(2,100)
DIMENSION OI(2,2),Z(100,2),MEAN(2),S(2,2),X(2),HOV(2,2),
ZZ(100,2),P1(2,2),SCALE(2),TOT1(2,2),TOT2(2,2),
U(80),V(80),Y(100,2)
DIMENSION D(160),GR(4),IND(8),ARPS(1),PMAS(1)
DIMENSION ASM(2,2),ASH1(2,2),ASM2(2,2),CMSE1(2,2),CMSE2(2,2),
CMSE3(2,2),MSE1(2,2),MSE2(2,2),AVCV(2,2),CVMS(2,2),EA(2,2),EB(2,2)
DIMENSION SUM(2,2),SUM1(2,2),SUM2(2,2),DD(2,2)
DIMENSION MD(2,2),M1(2,2),M2(2,2)
DIMENSION A(2,2),B(2,2),P(2,2),F(2,2),SH(2,2),SM1(2,2),SM2(2,2),
IC(2,2),CC(2,2),AA(2,2),BB(2,2),E(2,2),G1(2,2),G2(2,2)
DIMENSION CO(2,2),C1(2,2),C2(2,2)

G=2
IND(1)=80
IND(2)=1
IND(3)=1
IND(4)=0
IND(6)=5
IND(7)=1
IND(8)=6
READ(S,/) NX,SEED

READ(S,1) (MEAN(I),I=1,G)
READ(S,1) (SCALE(I),I=1,G)
FORMAT(2I10)

DO 10 I=1,G

READ(S,2) (OI(I,J),J=1,G)

READ(S,2) (P1(I,J),J=1,G)

FORMAT(2F10,3)

READ(S,/) (S(I,J),J=1,G)

READ(S,/) (ASM(I,J),J=1,G)

READ(S,/) (ASH1(I,J),J=1,G)

READ(S,/) (ASM2(I,J),J=1,G)

READ(S,/) (HOV(I,J),J=1,G)

CONTINUE

XN=NX

DO 15 I=1,2

DO 15 J=1,2

TOT1(I,J)=0

TOT2(I,J)=0

CMSE1(I,J)=0

CMSE2(I,J)=0

CMSE3(I,J)=0

MSE1(I,J)=0

MSE2(I,J)=0

SUM(I,J)=0

SUM1(I,J)=0

SUM2(I,J)=0

C 00000001

C 00000002

START OF SEGMENT 002

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:0

C 002:0000:5

C 002:0002:0

C 002:0003:0

C 002:0004:1

C 002:0005:2

C 002:0006:4

C 002:0007:5

C 002:0009:1

FIB IS 0006 LONG

C 002:0013:0

C 002:001F:2

C 002:002C:2

C 002:002C:2

C 002:002D:0

C 002:0038:2

C 002:0049:2

C 002:0049:2

C 002:0058:0

C 002:0067:0

C 002:0076:0

C 002:0085:0

C 002:0094:0

C 002:0096:1

C 002:0097:0

C 002:0098:0

C 002:0099:0

C 002:009B:4

C 002:009E:2

C 002:00A1:0

C 002:00A3:4

C 002:00A6:2

C 002:00A9:0

C 002:00AB:4

C 002:00AE:2

C 002:00B1:0

```

DO 5 I=1,G
ZZ(1,1)=RNMR(SEED)
5 DD(2,1)=ZZ(1,1)
DO 6 J=2,20
DO 102 K=1,G
102 X(K)=RNMR(SEED)
DO 111 I=1,G
DD(1,1)=0
DO 111 K=1,G
111 DD(1,1)=DD(1,1)+S(I,K)*X(K)
DO 8 I=1,G
ZZ(J,1)=OI(I,1)*ZZ(J-1,1)+OI(I,2)*ZZ(J-1,2)+DD(2,1)-PI(I,1)*
DD(1,1)-PI(I,2)*DD(1,2)
DD(2,1)=DD(1,1)
8 CONTINUE
6 CONTINUE
DO 7 I=1,G
7 ZZ(1,1)=ZZ(20,1)
DO 50 JJ=1,NX
DO 11 J=2,80
DO 103 K=1,G
103 X(K)=RNMR(SEED)
DO 112 I=1,G
DD(1,1)=0
DO 112 K=1,G
112 DD(1,1)=DD(1,1)+S(I,K)*X(K)
DO 12 I=1,G
ZZ(J,1)=OI(I,1)*ZZ(J-1,1)+OI(I,2)*ZZ(J-1,2)+DD(2,1)-PI(I,1)*
DD(1,1)-PI(I,2)*DD(1,2)
Y(J,1)=(ZZ(J,1)+MEAN(I))/SCALE(I)
DD(2,1)=DD(1,1)
12 CONTINUE
11 CONTINUE
BT(1,1)=ZZ(1,1)
BT(1,2)=ZZ(1,2)
DO 52 J=2,80
DO 51 I=1,G
151 BT(J,1)=ZZ(J,1)-OI(I,1)*ZZ(J-1,1)-OI(I,2)*ZZ(J-1,2)+PI(I,1)*
BT(J-1,1)+PI(I,2)*BT(J-1,2)
51 CONTINUE
52 CONTINUE
DO 55 J=1,80
DO 55 I=1,G
55 T(I,J)=BT(J,I)
DO 53 L=1,G
DO 53 I=1,G
MLT(L,I)=0
DO 54 J=1,80
54 MLT(L,1)=MLT(L,I)+T(L,J)*BT(J,I)
R(L,I)=MLT(L,I)/80.
53 CONTINUE
DO 21 J=1,80
U(J)=Y(J,1)
21 V(J)=Y(J,2)
DO 23 I=1,G
IND(5)=150
IF (I.EQ.2) GO TO 26
ARPS(1)=.7
PHAS(1)=.6
CALL FTXL(U,IND,ARPS(1),PHAS(1),PMAC,HNW,GR,D,IFR)
A(1,1)=ARPS(1)

```

```

C 002:0088:0
C 002:0089:0
C 002:008C:1
C 002:00C1:5
C 002:00C3:0
C 002:00C4:0
C 002:00C8:4
C 002:00CA:0
C 002:00CC:1
C 002:00CD:0
C 002:00D7:1
C 002:00D8:0
C 002:00E1:5
C 002:00E4:5
C 002:00E8:2
C 002:00EA:3
C 002:00EC:4
C 002:00EE:0
C 002:00F3:5
C 002:00F5:0
C 002:00F6:0
C 002:00F7:0
C 002:00FB:4
C 002:00FD:0
C 002:00FF:1
C 002:0100:0
C 002:010A:1
C 002:010B:0
C 002:0114:5
C 002:0117:5
C 002:011E:5
C 002:0122:2
C 002:0124:3
C 002:0126:4
C 002:0128:0
C 002:0129:4
C 002:0128:0
C 002:012C:0
C 002:0136:3
C 002:013A:4
C 002:013C:5
C 002:013F:0
C 002:0140:0
C 002:0141:0
C 002:014A:0
C 002:0148:0
C 002:014C:0
C 002:014E:4
C 002:0150:0
C 002:0159:4
C 002:015E:5
C 002:0163:1
C 002:0164:0
C 002:0166:2
C 002:0168:0
C 002:016C:0
C 002:016D:2
C 002:016E:4
C 002:0171:1
C 002:0173:1
C 002:0179:4

```


41
50

002:017C:5
002:017D:5
002:017E:5
002:0181:1
002:0187:4
002:0189:1
002:018A:4
002:018C:5
002:018E:0
002:018F:0
002:0191:4
002:0198:4
002:019A:0
002:019B:0
002:01A6:2
002:01A9:0
002:01AB:0
002:01AD:4
002:01AF:0
002:01B0:0
002:01B1:0
002:01BD:5
002:01CD:1
002:01D2:2
002:01D3:5
002:01DE:3
002:01E1:0
002:01F3:0
002:01F5:3
002:01F7:4
002:01F9:1
002:01FE:J
002:01FD:4
002:01FE:5
002:01F1:4
002:01F3:4
002:01F6:1
002:01F8:2
002:01FA:1
002:01FB:0
002:01FC:0
002:0201:1
002:0206:2
002:020F:1
002:0216:0
002:021B:1
002:0223:0
002:022A:5
002:0232:4
002:023A:3
002:023F:4
002:0244:5
002:024E:2
002:0250:3
002:0252:0
002:0253:0
002:0258:1
002:025D:2
002:0260:4
002:0264:4

```

CVMS(1,J)=CVMS(1,J)/XN
CMSE1(1,J)=CMSE1(1,J)/XN
CMSE2(1,J)=CMSE2(1,J)/XN
CMSE3(1,J)=CMSE3(1,J)/XN
MU(1,J)=SUM(1,J)/XN
M1(1,J)=SUM1(1,J)/XN
M2(1,J)=SUM2(1,J)/XN
44  WRITE(16,80)
80  FORMAT(/,10X,"ESTIMATED M0")
    CALL DTPUT(M0)
    WRITE(16,81)
81  FORMAT("0",19X,"M1")
    CALL DTPUT(M1)
    WRITE(16,82)
82  FORMAT("0",19X,"M2")
    CALL DTPUT(M2)
    WRITE(16,89)
89  FORMAT(/,5X,"MEAN SQUARE ERROR M0")
    CALL DTPUT(CMSE1)
    WRITE(16,90)
90  FORMAT("0",22X,"M1")
    CALL DTPUT(CMSE2)
    WRITE(16,91)
91  FORMAT("0",22X,"M2")
    CALL DTPUT(CMSE3)
    WRITE(16,83)
83  FORMAT(/,10X,"ESTIMATED C")
    CALL DTPUT(CA)
    WRITE(16,84)
84  FORMAT("0",19X,"E")
    CALL DTPUT(EB)
    WRITE(16,85)
85  FORMAT(/,5X,"MEAN SQUARE ERROR C")
    CALL DTPUT(MSE1)
    WRITE(16,86)
86  FORMAT("0",22X,"E")
    CALL DTPUT(MSE2)
    WRITE(16,87)
87  FORMAT(/,10X,"ESTIMATED COVARIANCE")
    CALL DTPUT(CVCOV)
    WRITE(16,88)
88  FORMAT("0",5X,"MEAN SQUARE ERROR OF COVARIANCE")
    CALL DTPUT(CVMS)
    STOP
    END

```

```

002:02C7:0  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:00B5
002:02C8:2  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0076
002:02C9:4  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0067
002:02CB:0  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:005A
002:02CC:2  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0049
002:02CD:4  IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0009

```

SEGMENT 002 IS 02E6 LONG

```

C 002:02C7:2
C 002:026A:4
C 002:026E:0
C 002:0271:2
C 002:0274:4
C 002:0279:5
C 002:027F:0
C 002:0288:3
C 002:028C:2
C 002:028C:2
C 002:028D:4
C 002:0291:2
C 002:0291:2
C 002:0292:4
C 002:0296:2
C 002:0296:2
C 002:0297:4
C 002:0298:2
C 002:0298:2
C 002:029C:4
C 002:02A0:2
C 002:02A1:4
C 002:02A5:2
C 002:02A5:2
C 002:02A6:4
C 002:02AA:2
C 002:02AA:4
C 002:02AF:2
C 002:02AF:2
C 002:02B0:4
C 002:02B4:2
C 002:02B4:2
C 002:02B5:4
C 002:02B9:2
C 002:02B9:2
C 002:02BA:4
C 002:02BE:2
C 002:02BF:2
C 002:02BF:4
C 002:02C3:2
C 002:02C3:2
C 002:02C4:4
C 002:02C5:3

```



```

12  SUBROUTINE TRAN(A,AA)
    DIMENSION A(2,2),AA(2,2)
    DO 12 I=1,2
    DO 12 J=1,2
    AA(I,J)=A(J,I)
    RETURN
    END

```

```

START OF SEGMENT 00A
C 00A:0000:0
C 00A:0000:0
C 00A:0000:0
C 00A:0001:0
C 00A:0002:0
C 00A:000C:0
C 00A:000C:3
SEGMENT 00A IS 0011 LONG

```

```

20  SUBROUTINE OUTPUT(A)
101 DIMENSION A(2,2)
    DO 20 I=1,2
    WRITE(16,101) (A(I,J),J=1,2)
    FORMAT(X,2E15,7)
    RETURN
    END

```

```

START OF SEGMENT 00C
C 00C:0000:0
C 00C:0000:0
C 00C:0000:0
C 00C:0001:0
C 00C:0010:3
C 00C:0010:3
C 00C:0011:0
SEGMENT 00C IS 0017 LONG

```

Example

Input:	Mean	0	0	
	Scale	1	1	
	OI	.900	.100] 1st variable
	PI	.400	.100	
	S	.36315, .547		
	ASM	.99, .3		
	ASML	.7264, .1162		
	ASM2	.6537996, .10458		
	HOV	.432, .2324		
	OI	.100	.800] 2nd variable
	PI	.100	.200	
	S	-.313885, .633		
	ASM	.3, .999		
	ASML	.13944, .6999		
	ASM2	.11155, .56		
	HOV	.2324, .5		

Continue Example OutputESTIMATED μ_0

.1030000E+01	.2733108E+00
.2733108E+00	.1900000E+01

 μ_1

.7177833E+00	.1052457E+00
.1157202E+00	.6790808E+00

 μ_2

.0395536E+00	.9086183E-01
.9041459E-01	.5557084E+00

MEAN SQUARE ERROR μ_0

.1000000E-05	.6283344E-02
.6283344E-02	.1000000E-05

 μ_1

.1325418E-01	.1164484E-02
.2213466E-02	.1476698E-01

 μ_2

.2644683E-01	.8037311E-03
.1189375E-02	.2336513E-01

ESTIMATED C

.8794567E+00	0.
0.	.5056161E+00

 σ

.3692978E+00	0.
0.	.2530591E+00

MEAN SQUARE ERROR C

.9464041E-02	.1000000E-01
.1000000E-01	.1463797E-01

 σ

.2135154E-01	.1000000E-01
.1001000E-01	.3420530E-01

ESTIMATED COVARIANCE

.4387918E+00	.2182003E+00
.2182003E+00	.4869830E+00

MEAN SQUARE ERROR OF COVARIANCE

.4699293E-02	.2979122E-02
.2979122E-02	.6141037E-02

Appendix B

Computer Program in

O'Connell's Method


```
FILE 16(KIND=PRINTER)
FILE 6(KIND=DISK)
```

```

INTEGER SEED
REAL MA,MO,M1,M2,MSE1,MSE2
DOUBLE PRECISION RHDG,CHDG
COMMON/HDG/RHDG(10),CHDG(10)
COMMON/COM1/N,NYR,NYG,NPC,MARKOV,AB(10),SCALE(10),NTRACE,IW
,IN,IA
COMMON/COM2/Y(150,10),XX(10,150),YPC(10,150),X(150,10),V(10),
PCA(10,10),NCORE,NX
COMMON/COM3/T(2,10),U(108),COR(3,10,10),WXM(10),SDY1(10),XMNY1(10)
,YIC(10)
COMMON/COMB/A1(10),SLN,WSD(10),A2(10),ISKEW
COMMON/MWNAT/CU(5),DV(5),CSNB(5),CSNM(5),CSLAG(5),STOR(5,1),NAT
COMMON/COH1/Q(1020)
COMMON/COH2/LENGTH,INT,IOPT5
COMMON/COH3/R(20,20),NSTOP(20),JMAX,NN(20)
COMMON/COH4/ARM(20),SIG(20),BT3
COMMON/PLOT/GRAPH(200),IOPT4
DATA N/2/,NYR/80/,NYG/80/,NPC/0/,IPC/0/,NTRACE/7/,ISKEW/1/,NX/2/,
NAT/0/,IONE/1/,IW/2/
INT=7
LENGTH=NYR
IOPT5=0
IOPT4=0
DIMENSION SS(2,2),TT(2,2),UU(2,2),BB(2,2)
DIMENSION OI(2,2),Z(100,2),MEAN(2),S(2,2),XB(2),
PI(2,2),ZZ(100,2),DD(2,2)
DIMENSION AMO(2,2),AM1(2,2),AM2(2,2),TEMP1(2,2),TEMP2(2,2),
AMTR1(2,2),AMTR2(2,2),SECOND(2,2),THIRD(2,2),FORTH(2,2),AINV1(2,2)
DIMENSION BBTR(2,2),BBINV(2,2),CC(2,2),TEMP3(2,2),
WKAREA(10),AAMO(6,6),AAM1(6,6),AAM2(6,6)
DIMENSION EE(2,2),BINV(2,2),DX(2,2),DXB(2,2),VAX(2)
DIMENSION CO(2,2),C1(2,2),C2(2,2)
DIMENSION ASM(2,2),ASM1(2,2),ASM2(2,2),CMSE1(2,2),CMSE2(2,2),
CMSE3(2,2),MSE1(2,2),MSE2(2,2),AVCV(2,2),CVMS(2,2),EA(2,2),ER(2,2)
DIMENSION SUM(2,2),SUM1(2,2),SUM2(2,2)
DIMENSION MO(2,2),M1(2,2),M2(2,2)
DIMENSION STD(2),TOT1(2,2),TOT2(2,2)
DIMENSION A(2,2),B(2,2),H(2,2),F(2,2),SM(2,2),SM1(2,2),SM2(2,2),
C(2,2),CX(2,2),AA(2,2),E(2,2),G1(2,2),G2(2,2)
DIMENSION HOV(2,2),MA(2)
G=2
IDGT=4
READ(5,/) NNX,SEED

```

```

READ(5,1) (MEAN(I),I=1,G)
READ(5,1) (SCALE(I),I=1,G)
FORMAT(2I10)
DO 10 I=1,G
  READ(5,2) (OI(I,J),J=1,G)
  READ(5,2) (PI(I,J),J=1,G)
  READ(5,1) (S(I,J),J=1,G)
  READ(5,1) (ASM(I,J),J=1,G)
  READ(5,1) (ASM1(I,J),J=1,G)

```

```

C 00000001
C 00000002
START OF SEGMENT 002
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0000:0
C 002:0001:1
C 002:0002:3
C 002:0003:4
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0004:5
C 002:0005:4
C 002:0006:3
FIB IS 0006 LONG
C 002:0010:0
C 002:001C:2
C 002:0029:2
C 002:0029:2
C 002:002A:0
C 002:0038:2
C 002:0046:2
C 002:0055:0
C 002:0064:0

```

```

10 READ(5,/) (ASH2(I,J),J=1,G)
   READ(5,/) (HOV(I,J),J=1,G)
   CONTINUE
2   FORMAT(2F10.3)
   XN=NNX
   DO 5 I=1,G
     ZZ(1,I)=RNMR(SEED)
5     DD(2,I)=ZZ(1,I)
     DO 6 J=2,20
       DO 102 K=1,G
102      XB(K)=RNMR(SEED)
        DO 111 I=1,G
          DD(1,I)=0
          DO 111 K=1,G
111         DD(1,I)=DD(1,I)+S(I,K)*XB(K)
          DO 8 I=1,G
            ZZ(J,I)=OI(I,1)*ZZ(J-1,1)+OI(I,2)*ZZ(J-1,2)+DD(2,I)-PI(I,1)*
            DD(1,1)-PI(I,2)*DD(1,2)
            DD(2,I)=DD(1,I)
8          CONTINUE
6          CONTINUE
          DO 7 I=1,G
7            ZZ(1,I)=ZZ(20,I)
            DO 50 JJ=1,NNX
              DO 69 I=1,G
                VA(I)=0
69                MA(I)=0
                DO 11 J=2,80
                  DO 103 K=1,G
103                 XB(K)=RNMR(SEED)
                    DO 112 I=1,G
                      DD(1,I)=0
                      DO 112 K=1,G
112                     DD(1,I)=DD(1,I)+S(I,K)*XB(K)
                        DO 12 I=1,G
                          ZZ(J,I)=OI(I,1)*ZZ(J-1,1)+OI(I,2)*ZZ(J-1,2)+DD(2,I)-PI(I,1)*
                          DD(1,1)-PI(I,2)*DD(1,2)
                          Y(J,I)=(ZZ(J,I)+MEAN(I))/SCALE(I)
                          Q(J)=Y(J,I)
                          DD(2,I)=DD(1,I)
12                         CONTINUE
11                        CONTINUE
                          DO 66 I=1,G
                            DO 68 J=1,80
                              MA(I)=MA(I)+Y(J,I)
68                             VA(I)=VA(I)+Y(J,I)**2
66                             STD(I)=SQRT((VA(I)-MA(I)**2./80.)/79.)
                              DX(1,I)=STD(I)
                              DX(2,I)=STD(2)
                              CALL MSTAT
                              DO 60 J=5,6
                                DO 60 K=1,G
60                                 AAM2(J,K)=COR(1,J,K)
                                    DO 63 J=4,5
                                      DO 63 K=1,G
                                        AAM0(J,K)=COR(2,J,K)
63                                     AAM1(J,K)=COR(3,J,K)
                                        AM0(1,1)=AAM0(4,1)
                                        AM0(1,2)=AAM0(4,2)
                                        AM0(2,1)=AM0(1,2)
                                        AM0(2,2)=AAM0(5,2)

```

```

C 002:0073:0
C 002:0082:0
C 002:0091:0
C 002:0093:1
C 002:0093:1
C 002:0094:0
C 002:0095:0
C 002:0098:1
C 002:0090:5
C 002:009F:0
C 002:00A0:0
C 002:00A4:4
C 002:00A6:0
C 002:00A8:1
C 002:00A9:0
C 002:00B3:1
C 002:00B4:0
C 002:00BD:5
C 002:00C0:5
C 002:00C4:2
C 002:00C6:3
C 002:00C8:4
C 002:00CA:0
C 002:00CF:5
C 002:00D1:0
C 002:00D2:0
C 002:00D3:3
C 002:00D7:1
C 002:00D8:0
C 002:00D9:0
C 002:00DD:4
C 002:00DF:0
C 002:00E1:1
C 002:00E2:0
C 002:00FC:1
C 002:00FD:0
C 002:00F6:5
C 002:00F9:5
C 002:0101:0
C 002:0104:3
C 002:0108:0
C 002:010A:1
C 002:010C:2
C 002:010D:0
C 002:010E:0
C 002:0112:0
C 002:0118:3
C 002:0120:1
C 002:0121:3
C 002:0123:0
C 002:0123:4
C 002:0125:0
C 002:0126:0
C 002:012F:3
C 002:0131:0
C 002:0132:0
C 002:0137:1
C 002:0140:4
C 002:0142:1
C 002:0143:5
C 002:0145:2

```



```

AM1(1,1)=AAM1(4,1)
AM1(1,2)=AAM1(4,2)
AM1(2,1)=AAM1(5,1)
AM1(2,2)=AAM1(5,2)
AM2(1,1)=AAM2(5,1)
AM2(1,2)=AAM2(5,2)
AM2(2,1)=AAM2(6,1)
AM2(2,2)=AAM2(6,2)
CALL LINV2F(AM1,2,2,AINV1,IDGT,WKAREA,IER)
CALL TRAN(AM2,AMTR2)
CALL TRAN(AM1,AMTR1)
CALL MULT(AM2,AINV1,TEMP1)
DO 67 I=1,G
DO 67 J=1,G
67 A(I,J)=TEMP1(I,J)
CALL MULT(TEMP1,AMTR1,SECOND)
CALL MULT(TEMP1,AM0,TEMP2)
CALL TRAN(TEMP1,TEMP3)
CALL MULT(TEMP2,TEMP3,THIRD)
CALL MULT(AM1,TEMP3,FORTH)
DO 61 J=1,G
DO 61 K=1,G
61 SS(J,K)=AM0(J,K)-SECOND(J,K)+THIRD(J,K)-FORTH(J,K)
TT(J,K)=TEMP2(J,K)-AM1(J,K)
DI=1
97 X1=TT(1,1)**2/DI
X2=TT(1,2)**2/DI
X3=TT(2,1)**2/DI
X4=TT(2,2)**2/DI
X5=TT(1,1)*TT(1,2)/DI
X6=TT(1,1)*TT(2,1)/DI
X7=TT(1,1)*TT(2,2)/DI
X8=TT(1,2)*TT(2,2)/DI
X9=TT(2,1)*TT(2,2)
X10=(TT(1,2)*TT(2,1)+TT(1,1)*TT(2,2))/DI
UU(1,1)=((SS(1,1)-SS(1,1)*X10+2*SS(1,2)*X5)*(X1*X9-X5*DI-X3*X5*DI)
-(SS(1,1)*X9-SS(2,2)*X5*DI)*(X1-X1*X10+2.*X5*X6))/
.(FLOAT(1-X10+X2-X2*X10+2.*X7*X8)*
.(X1*X9-X5*DI-X3*X5*DI)-(X9+X2*X9-X4*X5*DI)*(X1-X1*X10
+2.*X5*X6))
UU(2,2)=((SS(1,1)-SS(1,1)*X10+2.*SS(1,2)*X5)-UU(1,1)*FLOAT(1.
-X10+X2-2.*X2*X10+2.*X7*X8))/(X1-X1*X10+2.*X5*X6)
UU(1,2)=(SS(1,2)-UU(2,2)*X6-UU(1,1)*X8)/FLOAT(1.-X10)
UU(2,1)=UU(1,2)
DET=UU(1,1)*UU(2,2)-UU(1,2)**2
DIF=DI-DET
IF (ABS(DI-DET).LT..001) GO TO 98
DI=(DI+DET)/2
GO TO 97
98 BB(1,1)=ABS(SQRT(UU(1,1)))
BB(1,2)=UU(1,2)/BB(1,1)
BB(2,2)=ABS(SQRT(UU(2,2)-BB(1,2)**2))
CALL MULT(0X,BB,0XB)
CALL TRAN(0XB,0XBTR)
CALL MULT(0XBTR,0XB,H)
CALL TRAN(BB,BBTR)
CALL LINV2F(BBTR,2,2,BBINV,IDGT,WKAREA,IER)
CALL MULT(TT,BBINV,CX)
CALL LINV2F(BB,2,2,BINV,IDGT,WKAREA,IER)
CALL MULT(CX,BINV,B)
56 DO 20 I=1,G

```

```

C 002:0147:0
C 002:0148:3
C 002:014A:1
C 002:014B:4
C 002:014D:2
C 002:014E:5
C 002:0150:3
C 002:0152:0
C 002:0153:4
C 002:0158:0
C 002:015A:0
C 002:015C:0
C 002:015E:4
C 002:0160:0
C 002:0161:0
C 002:016A:0
C 002:016C:4
C 002:016F:2
C 002:0171:2
C 002:0174:0
C 002:0176:4
C 002:0178:0
C 002:0179:0
C 002:0184:4
C 002:0190:0
C 002:0190:4
C 002:0192:3
C 002:0194:3
C 002:0196:2
C 002:0198:2
C 002:019A:4
C 002:019C:5
C 002:019F:1
C 002:01A1:4
C 002:01A3:3
C 002:01A7:2
C 002:01AE:5
C 002:01B4:5
C 002:01B8:3
C 002:01BF:2
C 002:01C2:1
C 002:01C7:0
C 002:01CE:3
C 002:01D3:3
C 002:01D5:0
C 002:01D8:0
C 002:01D9:2
C 002:01DC:4
C 002:01DE:3
C 002:01DF:0
C 002:01E1:2
C 002:01F3:4
C 002:01E7:3
C 002:01EA:1
C 002:01EC:1
C 002:01FE:5
C 002:01F0:5
C 002:01F5:1
C 002:01F7:5
C 002:01FC:1
C 002:01FE:5

```

```

DO 20 J=1,G
SM1(I,J)=0
20 SM(I,J)=0
DO 71 I=1,G
DO 71 J=1,G
71 C(I,J)=A(I,J)-B(I,J)
CALL MULT(C,H,E)
CALL TRAN(C,CC)
CALL MULT(E,CC,F)
DO 72 I=1,G
DO 72 J=1,G
DO 73 K=1,29
SM1(I,J)=SM1(I,J)+A(I,J)**K*F(I,J)*A(I,J)**(K-1)
73 SM(I,J)=SM(I,J)+A(I,J)**(K-1)*F(I,J)*A(I,J)**(K-1)
SM1(I,J)=SM1(I,J)+E(I,J)
72 SM(I,J)=SM(I,J)+H(I,J)
CALL MULT(A,SM1,SM2)
TEMP=SQRT(SM(1,1)*SM(2,2))
CO(1,1)=SM(1,1)/SM(1,1)
CO(2,2)=SM(2,2)/SM(2,2)
CO(1,2)=SM(1,2)/TEMP
CO(2,1)=CO(1,2)
C1(1,1)=SM1(1,1)/SM(1,1)
C1(2,2)=SM1(2,2)/SM(2,2)
C1(1,2)=SM1(1,2)/TEMP
C1(2,1)=SM1(2,1)/TEMP
C2(1,1)=SM2(1,1)/SM(1,1)
C2(2,2)=SM2(2,2)/SM(2,2)
C2(1,2)=SM2(1,2)/TEMP
C2(2,1)=SM2(2,1)/TEMP
DO 41 I=1,G
DO 41 J=1,G
TOT1(I,J)=TOT1(I,J)+A(I,J)
TOT2(I,J)=TOT2(I,J)+B(I,J)
MSE1(I,J)=MSE1(I,J)+(A(I,J)-OI(I,J))**2
MSE2(I,J)=MSE2(I,J)+(B(I,J)-PI(I,J))**2
AVCV(I,J)=AVCV(I,J)+H(I,J)
CVMS(I,J)=CVMS(I,J)+(H(I,J)-HOV(I,J))**2
CMSE1(I,J)=CMSE1(I,J)+(CO(I,J)-ASM(I,J))**2
CMSE2(I,J)=CMSE2(I,J)+(C1(I,J)-ASM1(I,J))**2
CMSE3(I,J)=CMSE3(I,J)+(C2(I,J)-ASM2(I,J))**2
SUM(I,J)=SUM(I,J)+CO(I,J)
SUM1(I,J)=SUM1(I,J)+C1(I,J)
41 SUM2(I,J)=SUM2(I,J)+C2(I,J)
50 CONTINUE
DO 44 I=1,G
DO 44 J=1,G
EA(I,J)=TOT1(I,J)/XN
EB(I,J)=TOT2(I,J)/XN
MSE1(I,J)=MSE1(I,J)/XN
MSE2(I,J)=MSE2(I,J)/XN
AVCV(I,J)=AVCV(I,J)/XN
CVMS(I,J)=CVMS(I,J)/XN
CMSE1(I,J)=CMSE1(I,J)/XN
CMSE2(I,J)=CMSE2(I,J)/XN
CMSE3(I,J)=CMSE3(I,J)/XN
MO(I,J)=SUM(I,J)/XN
M1(I,J)=SUM1(I,J)/XN
44 M2(I,J)=SUM2(I,J)/XN
WRITE(16,80)
80 FORMAT(/,10X,"ESTIMATED MO")

```

```

C 002:0200:0
C 002:0201:0
C 002:0203:4
C 002:020A:4
C 002:020C:0
C 002:020D:0
C 002:0218:2
C 002:0218:0
C 002:021D:0
C 002:021F:4
C 002:0221:0
C 002:0222:0
C 002:0223:0
C 002:022F:5
C 002:023F:1
C 002:0244:2
C 002:024D:5
C 002:0250:3
C 002:0253:0
C 002:0255:0
C 002:0257:3
C 002:0259:4
C 002:025B:1
C 002:025D:1
C 002:025F:4
C 002:0261:5
C 002:0263:4
C 002:0265:4
C 002:0268:1
C 002:026A:2
C 002:026C:1
C 002:026D:0
C 002:026E:0
C 002:0273:1
C 002:0278:2
C 002:0280:1
C 002:0288:0
C 002:028D:1
C 002:0295:0
C 002:029C:5
C 002:02A4:4
C 002:02AC:3
C 002:02B1:4
C 002:02B6:5
C 002:02C0:2
C 002:02C2:3
C 002:02C4:0
C 002:02C5:0
C 002:02CA:1
C 002:02CF:2
C 002:02D2:4
C 002:02D6:0
C 002:02D9:2
C 002:02DC:4
C 002:02E0:0
C 002:02E3:2
C 002:02F6:4
C 002:02F8:5
C 002:02F1:0
C 002:02FA:3
C 002:02FE:2

```


	CALL OUTPUT(M0)	C 002:02FE:2
	WRITE(16,81)	C 002:02FF:4
81	FORMAT("0",19X,"M1")	C 002:0303:2
	CALL OUTPUT(M1)	C 002:0303:2
	WRITE(16,82)	C 002:0304:4
82	FORMAT("0",19X,"M2")	C 002:0308:2
	CALL OUTPUT(M2)	C 002:0308:2
	WRITE(16,89)	C 002:0309:4
89	FORMAT(//,5X,"MEAN SQUARE ERROR M0")	C 002:030D:2
	CALL OUTPUT(CMSE1)	C 002:030D:2
	WRITE(16,90)	C 002:030E:4
90	FORMAT("0",22X,"M1")	C 002:0312:2
	CALL OUTPUT(CMSE2)	C 002:0312:2
	WRITE(16,91)	C 002:0313:4
91	FORMAT("0",22X,"M2")	C 002:0317:2
	CALL OUTPUT(CMSE3)	C 002:0317:2
	WRITE(16,83)	C 002:0318:4
83	FORMAT(//,10X,"ESTIMATED C")	C 002:031C:2
	CALL OUTPUT(EA)	C 002:031C:2
	WRITE(16,84)	C 002:031D:4
84	FORMAT("0",19X,"E")	C 002:0321:2
	CALL OUTPUT(EB)	C 002:0321:2
	WRITE(16,85)	C 002:0322:4
85	FORMAT(//,5X,"MEAN SQUARE ERROR C")	C 002:0326:2
	CALL OUTPUT(MSE1)	C 002:0326:2
	WRITE(16,86)	C 002:0327:4
86	FORMAT("0",22X,"E")	C 002:0328:2
	CALL OUTPUT(MSE2)	C 002:0328:2
	WRITE(16,87)	C 002:032C:4
87	FORMAT(//,10X,"ESTIMATED COVARIANCE")	C 002:0330:2
	CALL OUTPUT(AVCV)	C 002:0330:2
	WRITE(16,88)	C 002:0331:4
88	FORMAT("0",5X,"MEAN SQUARE ERROR OF COVARIANCE")	C 002:0335:2
	CALL OUTPUT(CVMS)	C 002:0335:2
	STOP	C 002:0336:4
	END	C 002:0337:3
	002:0339:0 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0082	
	002:033A:12 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0073	
	002:033B:14 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0064	
	002:033D:10 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0055	
	002:033E:12 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0046	
	002:033F:14 IS THE LOCATION FOR EXCEPTIONAL ACTION ON THE I/O STATEMENT AT 002:0006	

SEGMENT 002 IS 035B LONG

Appendix CO'Connell's Table

$\varphi = 0.96$							
		$\tilde{E}[K]_n$			$\tilde{E}[\rho_1]_n$		
θ	$\rho_1 \backslash n$	25	50	100	25	50	100
0.92	0.058	0.654	0.661	0.657	-0.005	0.005	0.025
0.88	0.146	0.683	0.679	0.698	-0.005	0.038	0.080
0.84	0.250	0.681	0.704	0.724	0.060	0.083	0.132
0.80	0.357	0.693	0.737	0.767	0.069	0.160	0.227
0.76	0.457	0.718	0.752	0.781	0.099	0.196	0.264
0.72	0.545	0.735	0.776	0.811	0.144	0.259	0.366
0.68	0.620	0.756	0.796	0.827	0.219	0.310	0.444
0.64	0.683	0.763	0.818	0.843	0.243	0.382	0.504
0.60	0.734	0.791	0.823	0.855	0.300	0.424	0.567
0.56	0.776	0.803	0.825	0.861	0.353	0.475	0.621
0.52	0.810	0.801	0.852	0.866	0.377	0.565	0.647

Table (4.2) Values of $\tilde{E}[K]_n$ and $\tilde{E}[\rho_1]_n$ for selected values of φ and θ

$\varphi = 0.92$							
		$\tilde{E}[K]_n$			$\tilde{E}[\rho_1]_n$		
θ	$\rho_1 \backslash n$	25	50	100	25	50	100
0.88	0.049	0.657	0.664	0.654	0.001	0.012	0.028
0.84	0.114	0.687	0.680	0.686	0.005	0.046	0.079
0.80	0.189	0.686	0.705	0.709	0.072	0.093	0.123
0.76	0.269	0.699	0.735	0.745	0.082	0.160	0.208
0.72	0.349	0.725	0.751	0.756	0.116	0.199	0.240
0.68	0.426	0.740	0.782	0.783	0.169	0.269	0.332
0.64	0.496	0.772	0.783	0.800	0.218	0.309	0.390
0.60	0.560	0.773	0.796	0.803	0.273	0.364	0.437
0.56	0.616	0.774	0.820	0.825	0.285	0.432	0.516
0.52	0.665	0.794	0.825	0.828	0.335	0.467	0.532

Table (4.2) Values of $\tilde{E}[K]_n$ and $\tilde{E}[\rho_1]_n$ for selected values of φ and θ

$\varphi = 0.88$							
		$\tilde{E}[K]_n$			$\tilde{E}[\rho_1]_n$		
θ	$\rho_1 \backslash n$	25	50	100	25	50	100
0.84	0.046	0.659	0.665	0.651	0.005	0.014	0.029
0.80	0.102	0.689	0.678	0.676	0.012	0.050	0.078
0.76	0.176	0.688	0.703	0.696	0.080	0.097	0.117
0.72	0.233	0.703	0.728	0.728	0.091	0.158	0.196
0.68	0.302	0.729	0.745	0.737	0.126	0.198	0.225
0.64	0.370	0.742	0.772	0.764	0.176	0.262	0.309
0.60	0.435	0.771	0.773	0.778	0.223	0.302	0.363
0.56	0.495	0.771	0.788	0.782	0.278	0.356	0.407
0.52	0.550	0.774	0.808	0.802	0.291	0.416	0.481

Table (2.3) Values of $\tilde{E}[K]_n$ and $\tilde{E}[\rho_1]_n$ for selected values of φ and θ

$\varphi = 0.84$							
		$\tilde{E}[K]_n$			$\tilde{E}[\rho_1]_n$		
θ	$\rho_1 \backslash n$	25	50	100	25	50	100
0.80	0.044	0.660	0.663	0.647	0.007	0.016	0.029
0.76	0.096	0.690	0.676	0.668	0.016	0.052	0.077
0.72	0.154	0.688	0.699	0.687	0.086	0.099	0.113
0.68	0.214	0.705	0.720	0.714	0.098	0.156	0.189
0.64	0.277	0.725	0.746	0.733	0.142	0.203	0.237
0.60	0.325	0.753	0.749	0.747	0.180	0.245	0.290
0.56	0.409	0.753	0.766	0.753	0.240	0.298	0.332
0.52	0.454	0.761	0.784	0.773	0.254	0.355	0.406

Table (2.4) Values of $\tilde{E}[K]_n$ and $\tilde{E}[\rho_1]_n$ for selected values of φ and θ

$\varphi = 0.80$							
		$\tilde{E}[K]_n$			$\tilde{E}^A[p_1]_n$		
θ	$p_1 \backslash n$	25	50	100	25	50	100
0.75	0.055	0.666	0.657	0.649	-0.006	0.022	0.041
0.70	0.119	0.685	0.694	0.678	0.056	0.079	0.098
0.65	0.188	0.708	0.698	0.708	0.114	0.103	0.166
0.60	0.260	0.736	0.727	0.719	0.139	0.192	0.227
0.55	0.331	0.737	0.750	0.741	0.175	0.247	0.289
0.50	0.400	0.756	0.764	0.746	0.263	0.315	0.343

Table (3.5) Values of $\tilde{E}[K]_n$ and $\tilde{E}^A[p_1]_n$ for selected values of φ and θ

$\varphi = 0.75$							
		$\tilde{E}[K]_n$			$\tilde{E}^A[p_1]_n$		
θ	$p_1 \backslash n$	25	50	100	25	50	100
0.70	0.054	0.666	0.655	0.645	-0.004	0.023	0.041
0.65	0.115	0.684	0.689	0.671	0.059	0.079	0.096
0.60	0.179	0.707	0.691	0.698	0.117	0.102	0.161
0.55	0.246	0.733	0.719	0.707	0.141	0.189	0.219
0.50	0.313	0.734	0.740	0.728	0.177	0.240	0.277
0.45	0.377	0.751	0.754	0.732	0.264	0.307	0.329

Table (3.6) Values of $\tilde{E}[K]_n$ and $\tilde{E}^A[p_1]_n$ for selected values of φ and θ